DOCUMENT RESUME

ED 445 079		TM 031 756
AUTHOR	White, Amy E.	
TITLE	Result Generalizability and Detection of Discrepant Data Points: Illustrating the Jackknife Method.	
PUB DATE	2000-01-00	
NOTE	14p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Dallas, TX, January 27-29, 2000).	
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)	
EDRS PRICE	MF01/PC01 Plus Postage.	
DESCRIPTORS	*Data Analysis; Educational Research; *Estimation (Mathematics); *Generalization	
IDENTIFIERS	*Jackknifing Technique; *Research Replication	

ABSTRACT

The jackknife, as refined by J. Tukey (1958), is a valuable tool for the internal replication of a study. The jackknife statistic is particularly useful with small sample sizes. Large samples are labor intensive, and other methods better address this situation. The jackknife procedure involves the use of a single sample drawn from a normally distributed population. The jackknife procedure is a general method for reducing the bias in an estimator while providing a measure of the variance of the resulting estimator by sample reuse. The essence of the jackknife approach is to partition out the impact of a particular subset of the data on an estimate derived from the total sample. The method attempts to determine if any one case or group of cases exerts an inappropriate influence on the overall statistic of interest. To illustrate the value of the jackknife, an example is presented that uses actual educational research data. The study (B. White and L. Daniel, 1999) concerned career motivations of persons planning to teach. (Contains 4 tables and 13 references.) (SLD)

Result Generalizability and Detection of Discrepant Data Points

Result Generalizability and Detection of Discrepant Data Points:

Illustrating the Jackknife Method

Amy E. White

University of North Texas

Paper presented at the annual meeting of the Southwest Educational Research Association, Dallas, TX, January 27-29, 2000.

Result Generalizability and Detection of Discrepant Data Points:

Illustrating the Jackknife Method

The dictionary defines replicate as "to duplicate, copy, or repeat." Replication in research enhances credibility of the research process by eliminating that which sometimes contaminates research results; namely suspicion that the sample employed may be biased. Replication of research is the way scientists verify their results (Frymier, Gansneder & Robertson, 1989). Kessen (1960) termed replication as an elemental principle of competent research; however, replication of research studies is too often the exception rather than the rule. Examining psychological studies, Smith (1970) identified reasons why researchers fail to replicate: (a) lack of funds, (b) lack of time, (c) lack of availability of a comparable group of subjects, (d) development of new research interests, (e) desire to publish, (f) ego involvement with the data, and (g) the reluctance of some journals to publish replication studies. Frymier et. al (1989) stated, "Replication requires precise duplication of methods, instruments, and time lines if the concept is honored." (p. 3) Precision of this kind, though possible in the physical sciences, presents many obstacles in behavioral science research. Still, within education research, replication is essential to (a) check that the sample is not overly biased, and (b) verify that the results of the research are stable.

The purpose of all research is to generalize findings to the entire population, or to a population of interest. Therefore, it is imperative that research results, especially in education, be reported with confidence. The thinking researcher will take whatever steps necessary to produce credible results that can be generalized to the population of interest, including selection and reporting of the most significant data treatments and tests. One of the recognized triumvirates for reporting statistical results includes (a) statistical significance, (b) effect size, and (c) result replicability. There are philosophical issues even among the former. Larry Daniel's work concerning the misuse of statistical significance testing (Daniel, 1998) should be referenced for a clear discussion of these issues. However, it is the latter of the three, replicability that is so frequently ignored.

Problems with replication

The reticence of researchers to replicate their studies is understandable. Most statistical methods in use today were developed between 1800 and 1930 when computation was slow and expensive. Many researchers were trained under the assumption that replication was cost prohibitive and time consuming. A review of 20 standard textbooks on experimental design and methodology found that 12 did not list replication in their index, while 7 gave passing reference to the topic, and only Sidman (1960) treated the subject at length (Smith, 1970). However, with the advent of high-speed computers, and user-friendly statistical software, new methods for computing statistical analysis are fast and cost-efficient (Diaconis & Efron, 1983).

Computers have also allowed researchers to develop methods for another type of replication: internal replication. The more commonly known methods include cross-validation, the bootstrap by Efron, and the jackknife (Quenouille, 1949; Tukey, 1958). This paper concerns the latter: the jackknife

Using internal replication

The jackknife, as refined by Tukey (1958), is a valuable tool for internal replication of a study. The jackknife statistic is particularly useful with small sample sizes. Large samples are labor intensive, and other methods better address this situation.

The jackknife procedure involves the use of a single sample drawn from a normally distributed population (Schumacker, in press). The jackknife statistic is a general method for reducing the bias in an estimator while providing a measure of the variance of the resulting estimator by sample reuse. The essence of the jackknife approach is to partition out the impact of effect of a particular subset of the data on an estimate derived from the total sample (Crask & Perrault, 1977). The method attempts to determine if any one case or group of cases exerts an inappropriate influence on the overall statistic of interest (e.g., effect size, or R^2). With a single sample of data, the jackknife computes a new sample statistic for each sample size of K-1. The total number of psuedovalues (created by the jackknife) will equal the original sample size. The precise calculations for this statistic will later be discussed.

Data Example

To illustrate the value of the jackknife, a discussion using actual educational research data follows. The data (White & Daniel, 1999), used by permission, concerns career motivations of persons planning to teach. Initially, a regression analysis was completed on a randomly selected sample of White and Daniel's complete data set. The working sample can be viewed in Table 1. While the jackknife procedure can be used to

Insert Table 1 About Here

determine the stability of several statistics, the statistic of interest for the present study is the effect size, or R^2. It should be noted that the jackknife method is useful for determining the stability of several statistics (e.g., weights). However, it should also be noted that the effect sizes that are evaluated in statistical significance testing are more stable than the weights used to derive these effects (Thompson, 1994). Hence, the statistic of interest in all examples here mentioned is the effect size measure. An examination of Table 2 shows the R^2 for the initial regression was .391, an appreciable effect size. In order to determine the stability of this effect, the jackknife approach to internal replication was applied.

Insert Table 2 About Here

The calculation for the jackknife procedure is best addressed in Crask and Perrault's pivotal work (1977). The calculation is simply:

$$J_i = k(\Theta') - (k-1)\Theta_i$$

Where J_i = Statistic of interest computed without the first (2nd, 3rd, etc.) case
K = number of cases
Θ' = R^2 of regression all cases
Θ_i = R^2 of regression without the first (2nd, 3rd, etc) case
i = the given jackknife replication (i to K)

The procedure itself is simple. Each time the statistic of interest is computed, one case (or subset of cases) is eliminated. The statistic is computed K-1 times until each subset of data has been analyzed. The result is a pool of pseudovalues of the statistic from which the jackknife estimate can be calculated.

Insert Table 3 About Here

The remaining portion of the jackknife calculation may be employed to find the individual pseudovalues. Each of these values constitutes the error between the actual statistic of interest and the jackknifed estimate. The sum of the errors is divided by the number of cases, k ($\Sigma J_i /N$). Further, the researcher must obtain the mean and standard deviation of the jackknife calculations ($(k-1)\ \Theta_i$). In order to create a confidence interval, the critical value for the 95% confidence interval (1.96) is multiplied by the standard deviation. This value, added and subtracted from the mean, creates the confidence interval. The thoughtful researcher hopes the original statistic of interest falls between the confidence delimiters.

For the present example using the subset of White and Daniel's data (1999), the original $R^2 = .391$. Table 2 contains the recorded R^2 values for the subsequent regressions. The results for the example calculations can be found on table 3. The jackknifed value (.308)

Insert Table 4 About Here

resembles the original statistic (.391) and falls within the 95% confidence interval [4.74, -1.25]. The conclusion can be drawn that a certain stability exists within the research result estimator. A researcher could then report his/her results with an air of certainty and gain wider acceptance and credibility.

Conclusion

It is important to note that there is no substitute for actual, careful replication of a research study. Nothing adds more certainty to a result than its continued confirmation by further research studies. However, as previously mentioned, there are myriad reasons why actual replication is frequently improbable or impossible. So what is the answer? Researchers must begin to employ every tool at their disposal to add credibility to their work. In the field of education, the need for accuracy in reporting research is critical as these studies are so frequently used in decision-making. Why then do modern researchers neglect this vital need?

The jackknife method is labor intensive. It is only useful with small samples, and causes a researcher to regress many subsets of his/her data and calculate pseudovalues and confidence intervals. Occasionally, the confidence interval appears unrealistic because the researcher has forced the assumption of normality on data that may not be normally distributed. Internal replication is possible with the use of computers but still requires a measure of time and effort on the part of the researcher.

Even so, many tools are available to aid in the process of internal replication. Diaconis and Efron (1983) clearly illustrate the advantage of computer intensive research models. A researcher needing 40 regression analyses for a jackknife procedure need only "point and click" 40 times in a statistical package like SPSS. In light of the advantages computers bring to research, it is at least the duty of researchers to utilize them for the betterment of research. Internal replication studies of any kind, regardless of labor intensity are still easier, faster, and less expensive than the careful, accurate, and precise replication of a research study. And though internal replication is never as good as actual replication, it is still better than the replicability evidence that most researchers provide: nothing.

References

Ayabe, C.R. (1985). Multicrossvalidatioin and the jackknife in the estimation of shrinkage of the multiple coefficient of correlation. Educational and Psychological Measurement, 45, 445-451.

Campo, S. F. (1988). Alternative logics for estimating whether research results will generalize. Paper presented at the Annual Meeting of the Mid-South Educational Research Association, Louisville, KY. (ERIC Document Reproduction No. ED 315 439)

Crask, M. R., & Perreault, W. D. (1977). Validatioin of discriminant analysis in marketing research. Journal of Marketing Research, 14, 60-68.

Daniel, L.G. (1989). Use of the jackknife statistic to establish the external validity of discriminant analysis results. Paper presented at the annual meeting of the Southwest Educational Research Association, Houston, TX.

Daniel, L. G. (1998). Statistical significance testing: A historical overview of misuse and misinterpretation with implications for the editorial policies of educational journals. Research in the Schools, 5(2), 23-32.

Diaconis, P, & Efron, B. (1983). Computer-intensive methods in statistics. ScientificAmerican, 248(5), 116-130.

Frymier, J., Gansneder, B., & Robertson, N. (1990). Simultaneous replication: A technique for large-scale research. Paper presented at the annual meeting of the American Educational Research Association meeting, Boston, MA. (ERIC Document Reproduction Service No. ED 319 799)

Schumacker, R.E. (in press). Understanding statistical concepts using S-Plus simulation. New Jersey: Lawrence Erlbaum Associates, Inc.

Smith, N.C. (1970). Replication studies: A neglected aspect of psychological research. American Psychologist, 25, 970-975.

Thompson, B. (1994). The pivotal role of replication in psychological research: Empirically evaluating the replicability of sample results. Journal of Personality, 62(2), 157-176.

Tukey, J.W. (1958). Bias and confidence in not-quite large samples. Annals of Mathematical Statistics, 29, 614.

Quenouille, M.H. (1949). Approximate tests of correlation in time-series. Journal of the Royal Statistical Society, 11, 68-84.

White, B. L., & Daniel, L. G. (1999). Motivations for teaching as perceived by teacher education majors. The Professional Educator, 21(2), 39-51.

Table 1

Random Subset of White and Daniel's (1999) Data

Used by Permission

CASE	MAJOR	YRBIRTH	SERVINTR	SECURITY	CONTINUE	BENECOMP	STIMULAT	FILTER
1	ELEM ED	77	4.73	2.77	4.00	3.55	4.00	1
2	ENGLISH	75	3.47	2.46	3.40	2.73	3.00	1
3	HISTORY	76	3.87	1.62	3.80	2.36	3.56	1
4	ELEM ED	76	4.00	1.46	3.30	2.91	2.33	1
5	ELEM ED	77	3.73	2.23	2.50	2.00	2.44	1
6	SPEECH	76	4.20	2.23	4.20	3.18	4.00	1
7	ELEM ED	60	4.53	2.23	3.70	4.09	3.78	1
8	LIB SCI	34	3.80	3.00	2.50	2.18	4.22	1
9	ELEM ED	73	3.87	2.31	3.70	3.82	2.89	1
10	ELEM ED	72	3.53	1.85	4.00	2.27	2.67	1
11	ELEM ED	70	4.20	2.46	4.00	3.18	2.33	1
12	ELEM ED	76	3.53	2.31	3.60	3.18	3.11	1
13	SPEC ED	72	3.60	2.00	3.70	2.36	2.78	1
14	MUSICED	74	5.00	2.00	4.10	2.82	3.33	1
15	ELEM ED	74	3.33	2.85	2.20	2.73	4.44	1
16	SP PATH	74	4.00	1.85	3.90	2.64	3.33	1
17	ENGLISH	75	3.40	2.15	3.60	2.73	2.67	1
18	ELEM ED	67	4.80	1.31	4.40	2.73	2.67	1
19	ENGLISH	75	4.27	1.62	4.10	2.09	3.22	1
20	ELEM ED	77	4.33	2.15	3.90	3.00	3.33	1
21	SCIENCE	56	4.33	1.85	3.90	2.82	3.00	1
22	SPEC ED		4.67	3.77	3.40	4.55	3.89	1
23	SCIENCE	73	3.00	3.08	2.00	1.82	3.22	1
24	ELEM ED	76	4.67	2.00	4.30	3.36	3.67	1
25	MUSIC	77	4.00	1.46	4.20	3.45	3.22	1
26	ELEM ED	77	3.53	1.85	3.60	3.55	3.00	1
27	ELEM ED	72	3.80	1.31	4.10	3.45	3.00	1
28	HISTORY	75	3.73	1.54	4.40	1.91	2.22	1
29	ELEM ED	77	4.20	2.31	3.90	3.18	2.89	1
30	MUSIC	73	3.73	1.62	3.50	2.73	3.22	1
31	SP PATH	77	3.53	2.15	3.60	3.82	2.89	1
32	ELEM ED	71	4.53	3.38	2.60	3.00	4.22	1
33	ELEM ED	77	4.27	2.31	3.90	3.18	2.89	1
34	ELEM ED	69	4.80	1.77	4.50	3.55	3.78	1
35	SP PATH	77	4.87	2.77	2.00	3.18	3.67	1
36	ELEM ED	75	4.13	1.62	4.00	2.91	3.11	1
37	ELEM ED	77	4.47	1.92	4.10	3.09	2.89	1
38	SP PATH	75	3.40	3.31	2.50	3.09	3.89	1
39	BUS ED	72	4.20	4.08	4.30	2.27	2.56	1
40	ELEM ED	70	5.00	3.31	4.90	4.09	4.56	0

Table 2

Regression analysis with all cases present

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.626	.391	.322	.4245

Table 3

Results of K-1 regression analyses

Subject/Statistic	R^2	Pseudo-values
All cases	.391	
1.	.369	14.39
2.	.387	15.09
3.	.393	15.33
4.	.406	15.83
5.	.407	15.87
6.	.400	15.60
7.	.381	14.86
8.	.387	15.09
9.	.396	15.44
10.	.388	15.13
11.	.395	15.41
12.	.401	15.64
13.	.384	14.98
14.	.404	15.76
15.	.394	15.37
16.	.392	15.29
17.	.385	15.02
18.	.416	16.22
19.	.395	15.41
20.	.390	15.21
21.	.396	15.44
22.	.373	14.55
23.	.330	12.87
24.	.374	14.59
25.	.399	15.56
26.	.402	15.68
27.	.402	15.68
28.	.386	15.05
29.	.391	15.25
30.	.388	15.13
31.	.408	15.91
32.	.407	15.87
33.	.392	15.29
34.	.364	14.19
35.	.537	20.94
36.	.391	15.25
37.	.393	15.33
38.	.394	15.37
39.	.391	15.25
40.	.337	13.14

Table 4
Jackknife result for example data set

Results	
Σ of differences	12.76
Σ/K	.319
SD	1.21
Confidence Interval	[4.74, -1.25]
Conclusion	Stable Result

*The original effect (.391) falls within the 95% confidence interval

TM031756

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Result Generalizability and Detection of Discrepant Data Points: Illustrating the Jackknife Method

Author(s): Amy White

Corporate Source:

Publication Date:

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY Sample TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B
Level 1 ☑	Level 2A ☐	Level 2B ☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Sign here, → please

Signature: Amy E White	Printed Name/Position/Title: Amy E White
Organization/Address: University of North Texas	Telephone: 817 268 3596 FAX:
	E-Mail Address: amycox@yahoo.com Date: 1-29-2000

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, *or*, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2nd Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com